Facets of Assessment
Identification
Placement
Planning Specific Changes
Evaluating Intervention
Other Factors Shaping Assessment

Fundamental Concerns About Prevailing Practices
Problems of Interpretation
Biasing Factors
Ethics
Meeting Minimal Standards

Beyond Conventional Practices
Preassessment Interventions
Focal Point
Single- versus Multi-Stage Decision Making
Interventionist Assessment

Summing Up
Much time and effort is devoted to identifying individuals with learning problems and deciding what to do with them. As an integral part of the process, professionals give tests, carry out interviews, and make observations and ratings.

Attitudes toward prevailing assessment practices vary considerably. They are defended and cursed with equal vigor. Professionals are often confused by contradictory claims about the validity of a procedure. To the novice, it must be even more bewildering. Chapter 3 introduced the many purposes of assessment; this chapter is devoted to a more detailed discussion of this key component of intervention. Here we focus on each facet of assessment, emphasizing its uses, abuses, and limitations, and noting some contemporary trends.

**Facets of Assessment**

At the turn of the twentieth century, Alfred Binet and his co-workers undertook the task of developing assessment procedures to identify children unable to profit from regular instruction. Since that time, use of such procedures has been viewed as both necessary and desirable; this has led to a tremendous proliferation of assessment practices. However, as formal assessment practices became popular, widespread, and profitable, major concerns and controversies arose. This section details the four functions of assessment introduced in Chapter 3 (see Figure 3-3, p. 51).

**Identification**

Identifying individuals with learning problems is an intriguing detective job. Increasingly there are calls for large-scale searches for “underachievers,” “school failures,” “high challenges,” or “individuals at risk,” and of course for those with learning disabilities. As a result, screening designed to find and label those with learning problems, particularly young children, has become widely accepted.

The current ethos stresses that a person manifesting a problem ought to have professional help. Parents, teachers, doctors, legislators, and judges usually see it as their duty to refer the person to a professional helper. Such referrals often are based on informal assessment and thus are an informal type of screening.

Formal screening to identify persons who have problems or who are “at risk” is accomplished through individual or group assessment procedures. Group procedures are usually first-level screens that are expected to overidentify problems (that is, they identify many persons who do not have significant problems). Errors are supposed to be detected by follow-up assessments (see Feature 4-1).
Placement

The major practical objective of identifying problems is to correct them. Thus the ultimate value of finding problems depends on the nature of subsequent placement and corrective interventions.

As suggested in Chapter 3, placement considerations for individuals with learning problems involve understanding the problem's general nature and having information on available interventions. Once these matters are clarified, the placement process itself has three aspects:

- determining the general nature and form of interventions needed,
- deciding which of these to pursue, and
- from the service provider's perspective, deciding whom to serve.

Each of these is discussed below.

Nature and form of intervention needed. One of the most profound intervention decisions involves whether an individual should remain in the current setting or transfer to another. Such a decision is based on whether the person's problem is viewed as mild to moderate or severe and pervasive, and whether it is related to learning, behavior, or emotional functioning.

In some cases, adequate data for assessing severity and pervasiveness may come from information provided by those who have observed the individual's functioning in various settings (classroom, home, and recreation centers). When such information is not adequate, additional data can be gathered through psychoeducational testing, which may have been administered as part of the identification process.

When data indicate that a person is not making appropriate progress, what-

Feature 4-1 Concerns About Screening

Referrals based on informal assessment are commonplace. How people decide where to refer is an interesting—and sometimes controversial—matter. The literature that emphasizes a societal context suggests that referral is shaped by the status of both the referer and the person being referred. Socioeconomic status, race, education, age, and sex of both the referer and those referred can be major biasing factors. (We have more to say about factors that bias assessment later in this chapter.)

Even when referrals are based on formal screening instruments, there is reason for concern. We have mentioned the tendency to overidentify problems. In addition, screening procedures are often misused. For example, when data from first-level screening tests and rating scales are used to make tentative classifications and prescriptions, the labels and analyses are sometimes allowed to become definitive statements about an individual's status and needs. Given the low validity of first-level screens, especially those used with infants and primary grade children, this is unfortunate. At best, most screening procedures provide a preliminary indication that something may be significantly wrong. When diagnostic classifications and specific prescriptions are to be made, we need to use assessment procedures with greater validity.
ever the cause, the tendency is to consider use of special services and placements. Persons with severe and pervasive problems are often placed in special population settings such as special classrooms and institutions. Mild to moderate problems are supposed to be dealt with in normal population settings (that is, the mainstream)—either through modifying it somewhat or adding extra (ancillary) services or both. Placement decisions focus first on major intervention needs, then on which, if any, ancillary interventions seem indicated. In many cases, decisions about ancillary activity are best made after major interventions are given an adequate trial and found to be insufficient.

Ancillary interventions can involve

1. extra instruction such as tutoring
2. enrichment opportunities such as projects, learning by discovery, arts and crafts, and recreation
3. psychologically oriented treatments such as individual and family therapy
4. biologically oriented treatments such as medication

Assessment data for making decisions about extra instruction involve little more than an indication from a learner or teacher that current instructional time and resources are insufficient. Decisions to offer enrichment opportunities are not based on assessment data but on theory or belief about the value of such activity. The same is true of most psychological and biological interventions.
Client decision making. Recent years have seen increasing interest in consumer concerns. In Chapter 6 we highlight a variety of ethical and legal issues that have been raised about the role of clients in decision making.

Assessment can provide basic information for consumers about such matters as: What is available? What best meets the consumer’s needs in terms of location, cost, and type of intervention? and How can the consumer get admitted to a desired program? Unfortunately, even when a broad range of resources is available, as in large urban centers, systematic information is rarely presented to aid consumers in selecting services.

Professional and agency decisions. Once an application is made to a professional or agency, the focus shifts from the consumer to the service provider. Service providers use assessment data as aids in making placement decisions. Sometimes applicants are rejected because they do not fit established eligibility criteria. At other times, acceptable applicants must be rejected because too many have applied. Discussion of whom to select stresses data on type, severity, and pervasiveness of problem—and sometimes data on demographic factors.

Placement decisions involve institutional or individual values or both (Cronbach & Gleser, 1969). Data and criteria used in accepting and rejecting applicants usually reflect institutional values. For example, a professional or an agency may want only persons whose characteristics are viewed as maximizing the service’s efficacy, efficiency, or reputation. Thus selection may favor one sex, or individuals from certain backgrounds, or those with learning problems who are not hard to handle behaviorally.

It should be noted that the literature on placement accuracy is sparse. Moreover, validation studies are not reliable. If a person placed in a service does not improve, it may be unclear whether this was due to

1. selecting the wrong type of intervention,
2. the intractability of the problem, or
3. the way the intervention was implemented.

Given the likelihood of error, as well as other factors that can make a placement unsatisfactory, placement decisions must be reassessed regularly.

Planning Specific Changes

Prevailing assessment procedures for planning the specifics of an intervention reflect contrasting views toward remediation. They also reflect differing views about how to design instruction.

Remedial orientations shape assessment practices. Advocates of a particular orientation will usually argue for their procedures and sometimes argue against those reflecting other orientations. To understand why requires some knowledge of opposing views about what is to be remedied.

Essentially, current assessment practices have been dominated by the two remedial orientations discussed in Chapter 3—approaches emphasizing either
underlying problems or direct instruction of observable skills. As the emphasis on metacognitive (general learning) strategies has evolved, this has generated additional ideas about remedying learning problems. Each orientation defines remedial needs in different terms and thus specifies assessment differing in nature and scope.

(a) Underlying problems. Advocates of the underlying problem orientation (sometimes called the process or diagnostic-prescriptive model) postulate that learning problems result from disabilities related to specific areas of development. Thus they focus on assessing such concepts as perceptual, motor, linguistic, and memory functioning as a basis for planning treatment. They assume underlying problems must be corrected or compensatory strategies must be acquired before basics such as reading can be learned.

The ultimate value of the underlying problem orientation is yet to be determined. Available tests and research based on this orientation have been criticized severely (see reviews by Arter & Jenkins, 1979; Coles, 1978). In response, it has been argued that such criticism is not evidence that the basic premises of the underlying problem orientation are invalid (Colarusso, 1987; Snart, 1985). The poor validity of a given test says little about the theory upon which it is based. Recognizing this, some critics have directed their arguments against the theoretical assumptions of the approach (Carnine & Woodward, 1988). However, recent emphasis on cognitive and temperament components of learning problems has revitalized the theoretical underpinnings of this orientation and renewed interest in developing remedial and assessment approaches aimed at underlying problems.

(b) Observable problems. Those oriented to direct instruction of what is observable are concerned with specific knowledge and skills (including traditional readiness-prerequisite skills). They concentrate on what has not yet been learned by the individual. That is, they assess, teach, and reteach such basics as reading, language, mathematics, and social skills. Their skill-oriented assessments tend to use standardized achievement tests; unstructured, informal skill diagnostic tests; observation of daily performance; and criterion-referenced evaluations.

The observable nature of what is assessed has made this orientation attractive to those who write education agency guidelines. As a consequence, many such guidelines have been written in ways that shape the formulation and evaluation of programs in behavioral terms.

Advocates for focusing on underlying problems argue that such dysfunctions can interfere with learning skills and that, by ignoring underlying problems, direct instructors prevent some individuals from making progress. Motivation theorists, on the other hand, criticize both orientations for ignoring factors such as avoidance and anxiety. As to observable skills themselves, there has been controversy about what skills should be assessed and about whether all assessed skills need to be taught formally.

(c) Strategy problems. Those who pursue metacognitive (general learning) strategies emphasize the awareness and control learners can produce over the learning process. That is, the focus is on the ways an individual thinks about how to learn (one’s awareness of strategies for learning and one’s ability to direct one’s thinking about how to learn). Some interveners stress metacognitive skills (Brown
& Campione, 1986); others distinguish metacognitive knowledge from metacognitive experiences (Flavell, 1985; Garner, 1987).

The most used procedures in assessing metacognitive knowledge and cognitive monitoring are verbal interviews and error detection. Assessment procedures are chosen for their relevance to school learning and for their ease of use (Harris & Pressley, 1991; Wiener, 1986; Wong, 1991). Such procedures have been criticized both methodologically and conceptually (for example, problems of reliability and validity related to interviews, debates over what strategies can and should be taught).

The interest in cognitive components underlying learning problems goes beyond metacognitive knowledge and cognitive monitoring (Swanson, 1988). What is called the dynamic assessment movement (see Feature 4-2) is a variety of techniques for assessing cognitive skills and processes has been identified as appropriate for remedial intervention (Campion & Brown, 1987; Feuerstein, 1979).

We do not expect that issues surrounding current orientations to remediation will be resolved in the near future. Although specific instruments may be discredited, advocates of each orientation will continue to use assessment procedures that reflect their remedial approach and will continue to criticize each other. At the

---

**Feature 4-2 The Dynamic Assessment Movement**

The term *dynamic assessment* was coined by Feuerstein (1980). During the 1980s, the desire to use more interventionist approaches to assessment coalesced into what has come to be called the dynamic assessment movement. Dynamic assessment can be seen as a reaction to static (conventional psychometric) approaches to measuring intelligence. Static approaches are criticized for treating IQ as a trait rather than a score, thereby equating it with learning ability and ignoring the nature of cognitive development and functioning and the influence of handicapping conditions and cultural bias. Static approaches and techniques are also criticized as too limited for planning interventions.

The dynamic assessment movement wants to go beyond conventional psychometric techniques. It is interested in determining "the operation of basic psychological processes presumed responsible for acquisition of the information requested on standard tests" (Campion & Brown, 1987; p. 82).

Feuerstein and his colleagues (Feuerstein, Rand, Jensen, Kaniel, & Tzuriel, 1987) stress that their approach to dynamic assessment provides data not only on cognitive functioning but also on "structural cognitive modifiability ... the very structural nature of the cognitive processes that directly determine cognitive functioning in more than one area of mental activity" (pp. 42–43). Their intent is to assess the efficiency of specific cognitive processes, not just measure enhanced performance or the magnitude of response to instruction (see Haywood, Brown & Wingenfeld, 1990; Lidz, 1987).

Because dynamic assessment involves prompting and teaching, it is also used to identify intervention approaches that seem to work well for the individual. Thus the products of the assessment are seen as providing information on both what and how to teach. In order to gather such data, the process is designed and implemented as an intervention to improve performance. Improved performance is interpreted as an indication of the learner's real capability as well as the ability to change (learning potential).

Because the content focus of dynamic assessment is on underlying processes, critics of the movement have attacked the underlying problem orientation have attacked the movement. In doing so, they reiterate arguments that the measures and theoretical assumptions lack validity.
same time, advocates of environmental and transactional models can be expected to escalate their criticism and propose alternative assessment procedures.

**Matching instruction to the learner.** Differing views about how to design instruction for specific learners also lead to divergent assessment perspectives. For instance, concern has been raised that assessment for individualized, as contrasted with personalized, instruction results in an inadequate instructional design (Adelman & Taylor, 1983a; 1986).

*Individualization* typically emphasizes detecting a student’s deficiencies by monitoring daily performance and then modifying instruction to address the deficiencies. Approaches like dynamic assessment attempt to assess the best teaching approach for a given child. In most cases, however, a major shortcoming of assessment based on individualized instruction is that it overemphasizes developmental deficiencies and underemphasizes motivation—especially intrinsic motivation. This is not surprising, given how little systematic attention researchers and practitioners have paid to the concept of intrinsic motivation.

In contrast, the concept of personalization could broaden the focus of assessment. *Personalization* encompasses individualization by stressing the importance of designing intervention to match not only current learner capabilities but also levels of motivation, and especially intrinsic motivation. This latter emphasis is seen as critical, given the degree to which intrinsic motivation can profoundly affect current, as well as long-term, performance and learning (Adelman, 1978; Adelman & Taylor, 1990; Deci & Chandler, 1986). Thus a major implication of personalization is that systematic procedures are needed to address motivation.

All formal and informal procedures used to assess and prescribe specific treatment plans raise basic concerns (Salvia & Ysseldyke, 1991). Many experts suggest that among problem populations a person’s performance is often affected by low motivation or high anxiety. When this is the case, the findings are “contaminated.” Under such circumstances, it is impossible to know whether failure to demonstrate an ability represents a real deficiency; thus it is easy to misprescribe treatment.

It is easy, for example, to make the mistake of planning to teach skills that a person has already acquired—instead of helping the individual overcome psychological problems that interfere with performance. Also, concerns have been raised that comprehensive remedial plans need to go beyond what can be readily assessed by prevailing practices. Finally, it has been stressed that assessment may be viewed as only one facet of making intervention decisions.

**Evaluating Intervention**

It is not uncommon to hear professionals say, “if it works, use it!” Unfortunately, there is rarely adequate evidence about what really works in the long term. One reason that conflicting orientations exist is the difficulty of measuring what works and what doesn’t. A considerable amount of research has been reported; however, measurement and other research methodology problems have made it impossible to prove the worth of the programs studied (Tindal, 1985).
Nevertheless, assessment plays a major role in efforts to answer the basic question: Are interventions for learning problems effective? Although some may prefer to ignore this question, two facts make this impossible. One, evaluative research is essential to improve interventions. Two, this is an age of accountability, and evaluation is increasingly mandated by legislation, government regulations, and funding agencies. It is easy to mandate accountability. Unfortunately, such mandates ignore the fact that current evaluation practices are terribly inadequate (see Feature 4-3). Thus, while mandated evaluation goes on continually, comprehensive and valid evaluation is rare.

Chapter 12 discusses evaluating interventions. At this point, it is sufficient to indicate that evaluation involves determining the worth or value of something (Stake, 1967, 1976). In formal terms, we define evaluation as a systematic process designed to describe and judge the overall impact and value of an intervention for purposes of making decisions and advancing knowledge.

More specifically, the goals and objectives of evaluation include the following:

- to describe and judge an intervention’s (a) rationale, including assumptions and intentions, and (b) standards for making judgments
- to describe and judge an intervention’s (a) actual activity, including intended and unintended procedures and outcomes, and (b) costs (financial and negative effects)
- to make decisions about continuing, modifying, or stopping an intervention for an individual or for all those enrolled in a program

**Feature 4-3 Evaluation Problems and Issues**

Problems arise because of limitations in measurement capability. Unfortunately, many of the measurement instruments are not highly reliable or valid. Furthermore, because of accountability pressures there has been an overemphasis on measuring immediate behavioral outcomes. This is unfortunate, because many desirable program outcomes are not easily translated into immediate behaviors. (For example, self-concept, attitudes toward learning, problem-solving capabilities, creativity) and thus are deemphasized in planning and evaluating programs. Comprehensive evaluation requires a range of valid procedures, and development of such procedures requires considerable financial commitment. Because of the costs, evaluation remains a token item in most budgets.

Issues about evaluation arise because of different views about its appropriate focus (Is it sufficient to gather data on person variables? Should long-term outcomes be measured as well as immediate effects?) and about the best way to gather and interpret data (What specific measures and design should be used? Should there be emphasis on minimizing the negative effects of evaluation?). Because of differing views, issues arise over whose perspective should determine the evaluation focus, methods, and interpretation. Should the views of teachers, parents, students, researchers, or funding agencies prevail? Stated more boldly, whose biases or vested interests should prevail? (We have more to say about biases later in the chapter.)

Finally, we note that decisions as to what and how to evaluate are made by those administering or funding a program. Currently, there is little student or parent involvement in such decisions and, not surprisingly, there is little emphasis on client-consumer judgments of whether a program has value.
• to advance knowledge about interventions to improve (a) practices, (b) training, and (c) theory

The information needed to meet these purposes comes from comprehensive evaluations that include both immediate and long-term program data. Commonly, programs are evaluated using learner-centered paper-and-pencil tests of ability and performance, reports (including grades), and systematic interviews and observations of behavior. However, because such practices are so limited when it comes to assessing complex performance, a trend has evolved toward what is being called “authentic” assessment (Archibald & Newman, 1988; Linn, Baker, & Dunbar, 1991; Wiggins, 1989). The focus of this trend is on performance-based evaluation using such procedures as essays, open-ended responses, responses to computer simulations, interview data, and analyses of student journals and work that is accumulated over time in a portfolio.

**Other Factors Shaping Assessment**

In addition to having four major purposes, activity related to assessment occurs in phases, and differs in terms of focus and types of procedures used (see Table 4-1). Moreover, there are a variety of options in deciding what and how to assess.

For example, stimulus-and-response conditions may differ in terms of the number of variables assessed, their complexity, and whether they are simulated or natural. Variations also occur with respect to (a) how ambiguous and subjective the stimuli are, (b) how well standardized the administration procedures are, (c) how obtrusive the procedures are, and (d) how much they cause unintended reactions. There are also important considerations about similarities and differences between the assessor and the assessed (for example, in terms of race, cultural background, socioeconomic status, and gender).

Although we recognize that such variations in practice influence both the form of assessment activity and the findings, there is little agreement and considerable concern about their impact.

---

**Fundamental Concerns About Prevailing Practices**

Reviewers have been consistent in noting that most assessment practices for learning problems raise major concerns with respect to their validity as diagnostic and prescriptive tools (Coles, 1987; Fewell, 1991; Reynolds & Kaiser, 1990; Salvia & Ysseldyke, 1991). On a broader level, there has been concern over the way conventional assessment practices perpetuate an unsatisfactory status quo.

By understanding some fundamentals, one can appreciate the importance of new approaches that are designed to go beyond prevailing practices. In the preceding discussion, we alluded to a variety of concerns. This section takes a more systematic look at

- problems affecting interpretation of assessment data
- factors that bias assessment
- ethical concerns
### Table 4-1  Nature and Scope of Assessment Activity

I. Functions and purposes of assessment  
   A. Identification  
      1. Screening and referral  
      2. Diagnostic labeling  
      3. Nonpathological attributes  
   B. Selection/placement  
      1. Clarification of options  
      2. Client decisions about general changes in status  
      3. Professional and agency decisions to accept or reject applicants

C. Specific planning for change  
   1. Detailed objectives  
   2. Detailed procedures

D. Evaluation of interventions  
   1. Individual efficacy  
   2. Efficacy for all participants  
   3. Impact on society

II. Major phases related to assessment  
   A. Preparatory decisions about what is to be assessed (implicit or explicit rationale for assessment activity)  
   B. Description ("measurements" of specified variables and serendipitous data gathering, followed by analyses and descriptive summaries)

C. Judgments (interpretations)  
D. Communication and decision making with reference to assessment purposes

III. Focus of assessment  
   A. Focal point  
      1. Person(s) – individuals or groups of individuals  
      2. Environment(s)  
      3. Person–environment transactions  
   B. Nature of phenomena  
      1. Problematic–nonproblematic conditions  
      2. Observable–inferred  
      3. Proximal–distal  
      4. Historic–current–future expectations

C. Levels  
   1. Molecular–molar analyses of persons  
   2. Primary, secondary, tertiary contextual analysis  
   3. Transaction of person–environment

D. Areas or domains  
   1. Biological and psychological processes  
   2. Motor and verbal functioning  
   3. Physical environment  
   4. Social environment  
   5. Transaction of person–environment

IV. Types of procedures and instruments (standardized, semi-standardized, or unstandardized)  
   A. Interviews and written personal reports (responses to oral or written questions, inventories of items, etc.)  
   B. Observations  
   C. Verbal and performance measures (objective instruments such as achievement tests; projective instruments such as thematic pictures; instruments developed by teachers, psychologists, and MDs that have not been formally and technically standardized)

D. Biological tests (electrorecording devices, chemical analyses)  
E. Available records and data (analyses of current or cumulated records related to person, environment, transactions; analyses of natural performances and products, such as portfolio assessment)
Problems of Interpretation

Decisions about what data to gather are determined by views about what we want to assess. However, the actual data provide only a description of observed behavior. We do not directly observe intelligence or perception or attention deficits or minimal CNS dysfunctioning. Instead, we see responses to stimuli, or noncompliance with rules and directions (test answers or refusal to do a task), or reports of poor performance and misbehavior. Thus, concerns arise about the following:

Inadequate descriptions of a phenomenon
- Does the procedure provide information relevant to the types of interpretations and judgments we want to make?
- Does the procedure measure what it says it does?
- Only what it says it does?
- All of what it says it does?
- Does it provide new information?

Invalid interpretations and judgments of phenomena
- Are inferences justified?
- Are appropriate norms and standards available?

Invalid decisions
- How relevant are the data with respect to the decisions to be made?

Setting out to assess cognitive deficiencies, for example, we want to be able to
discuss findings in terms of such deficiencies and to judge the nature and scope of deficiencies based on appropriate norms and standards.

**Reliability and validity.** Deficiencies with regard to reliability and validity generate controversy not only about what findings mean but about the value of assessment per se. Obviously it is important for a procedure to be reliable (that is, to provide consistent and reproducible findings). Fortunately, reliability can be determined in a relatively technical and objective way. Unfortunately, many procedures used to assess learning problems are not highly reliable.

Validation of a procedure is not so straightforward; it requires a great deal of rational and subjective activity. As Cronbach (1970) stressed with respect to assessing constructs,

> construct validity is established through a long-continued interplay between observation, reasoning, and imagination.... The process of construct validation is the same as that by which scientific theories are developed. (p. 142)

Because determining a procedure's validity is difficult and costly, assessors find themselves having to use the best that is available, even though the best may not be very good. The inevitable result has been criticism of procedures and decisions based on them.

**Norms and standards.** Given the same set of findings, people still arrive at different conclusions. Why? Often because they use different standards in interpreting what they see. The term **standards** refers to values, theoretical ideas, and the empirical bases used to make judgments. The judgments often are about whether a problem exists, whether something is good or bad, and whether what was assessed is consistent with some theory.

**Norms** are empirical standards (not value or theoretical statements). That is, norms are a set of findings that can be used for purposes of comparison. (Are the current findings higher, lower, or the same as the earlier findings? How much higher or lower?) Formal norms are based on research and systematic observation. In practice, any set of findings (including nonsystematic observations over years of professional practice) might constitute the norms with which to compare assessment data.

Norms and standards are the major referents in interpreting assessment data. When norms are inadequate, or when there is no consensus regarding standards, interpretations become controversial. To understand this, it is important to appreciate the nature of the interpretive or judgmental process in assessment, which requires an appreciation of the way standards and norms are used.

After formal or informal norms are used, it is commonplace to apply value-based standards to make judgments. This can happen so quickly that it may not be apparent. For example, a score from a test or rating scale that falls outside some predetermined average may be quickly translated into a judgment that the performance indicated significant underachievement, developmental lag, or pathology. This is understandable, especially with tests of achievement and measures of development. However, it is important to note that a judgment has been made. The use of norms, of themselves, does not lead to a judgment.
All in all, controversy seems inevitable, given that assessors must rely on inadequate procedures and given the lack of consensus about standards (Buros, 1974; Coles, 1987; Schrag & Divoky, 1975). And the criticism is justified when professionals proceed in uncritical ways. As Buros (1974) noted, practitioners seem to have an unshakable will to believe the exaggerated claims of test authors and publishers. If these users were better informed regarding the merits and limitations of their testing instruments, they would probably be less happy and successful in their work. The test user who has faith—however unjustified—can speak with confidence in interpreting test results and making recommendations. The well-informed test user cannot do this; [this person] knows that the best of our tests are still highly fallible instruments which are extremely difficult to interpret with assurance in individual cases. Consequently, [the user] must interpret test results cautiously with so many reservations that others wonder whether [s]he really knows what [s]he is talking about. (p. xxxvii)

**Biasing Factors**

Potential for bias in assessment is great. Indeed, some degree of bias is inevitable. In this instance, we are not concerned with the prejudiced and stereotypic thinking of an individual, but with factors that affect large numbers of assessors and cause them to react in systematically biased ways.

Bias (that is, selectivity and distortion) affects both the descriptive and judgmental phases of assessment. Systematic bias stems from a weak knowledge base and from psychological and societal factors that affect the information available to the assessor and the way that information is processed.

**Weak knowledge base.** To comprehend why some degree of selectivity and distortion is inevitable, one need only recognize that assessors primarily look for what they have come to understand are the important aspects of phenomena. They cannot really afford to do otherwise. There is no time to describe everything, and random sampling of complex phenomena tends to produce random findings and little understanding of what is assessed. Thus, whether explicitly stated or not, each assessor is guided by some underlying rationale.

What does an assessor’s rationale consist of? Usually the assessor has some theory or model about what and how something is to be assessed. If there is widespread consensus about the assessor’s rationale, descriptions and judgments probably will not be seen as biased. That is, where there appears to be a strong base of knowledge related to a phenomenon and how to measure it (for example, X rays of broken bones), there will be little controversy over methods and conclusions. And, whatever selectivity and distortion are present will not be readily apparent.

In contrast, when theory and research are viewed as inconclusive, there will be diverse and competing models and methods. Assessors will differ, often markedly, in the rationales shaping their practices. They will try to draw on the available (albeit weak) knowledge base. However, because no model, orientation, or trend will be dominant, even professionals from the same discipline can be expected to be
guided by different ideas. Under such circumstances, assessors will look for different data, perhaps use different methods, and often arrive at different conclusions about the same phenomena. Moreover, it should not be too surprising that there will be a tendency to find and interpret data in ways that are consistent with one's theories.

A weak knowledge base means that research is only in its early stages and cannot resolve debates over which ideas and practices are correct. When there is a weak knowledge base, the factors underlying a practice's acceptance appear to be its compelling "logic" and intuitive appeal, or the proselytizing capability of its adherents. Many pioneering assessment practices that lack validity have become institutionalized through training programs, unsupported expert opinion, publishing company sales campaigns, and so forth. Such practices prevail until a strong body of research not only documents their inappropriateness but also is used effectively to undermine their institutionalized support. (However, even when there are data suggesting a given practice is ineffective or even harmful, the findings may not be accepted. Widespread testimony claiming a practice's validity can overwhelm arguments against its use.)

Currently, psychology and education are dominated by relatively weak theories and methods for measuring causes and correcting problems. This ensures ongoing controversy among those who hold competing assessment rationales. Selection and distortion continue to be inevitable.

Coexistence of contradictory perspectives does not mean that each enjoys equal status. A model may be in vogue, but vogues change. For example, for some time a pathological perspective (medical or disordered-person model) dominated psychoeducational assessment practices. Only in recent years has this view been challenged sufficiently to allow nonpathologically oriented assessment procedures to evolve.

With reference to the case of David, the weak knowledge base affected everyone involved in making the decisions. While the teacher knew David was not performing as well as others in the class, she didn't know why or what to do about it. She recalled some of the "symptoms" she had learned about in education courses and recognized these in David. This led her to feel justified in referring him for help.

The psychologist duly noted David's symptoms. She administered the standard assessment procedures used by the school district. Her diagnosis and recommendations were based on the view that persons who test significantly below grade level on an achievement test and who score at least average on an IQ test can be diagnosed as having a learning disability and might have a minimal brain dysfunction. Because David was a poor reader and did not do well on a test of visual perception, she thought he might be dyslexic.

David's parents felt confused about the problems. They recalled various TV programs and magazine articles dealing with learning disabilities, dyslexia, and brain damage. They were further mystified, yet strangely reassured, by the information provided by the school's "experts." In the end, they agreed to all the recommendations because they didn't think there was an alternative.

As for David, neither his parents nor the school personnel were certain about what to tell him or how to talk with him about the nature of the problem.
Psychological bias. In addition to bias that stems from the knowledge base, there is selectivity and distortion from psychological factors that affect the information available to the assessor and the way that information is processed. A few examples will illustrate this point.

As a general principle, the more complex and comprehensive the phenomena assessed, the more difficult it is to obtain an adequate sample of data. Although reasons for this may be strictly procedural and practical (lack of instruments, time, and so forth), sometimes the difficulty is due to an assessor's cognitive, perceptual, and motivational functioning. An example of one such psychological factor, of potentially great importance in understanding the bias in assessment, comes from attribution theory.

Those who assume the role of gathering data, especially professional assessors, can be described as observers of phenomena. Persons observed can be designated as actors. In discussing perceptions of the causes of behavior, attribution theorists have suggested a pronounced tendency for observers and actors to perceive the causes of an actor's behavior in different ways. Specifically, as hypothesized by Jones and Nisbett (1971), "there is a pervasive tendency for actors to attribute their actions to situational requirements, whereas observers tend to attribute the same actions to stable personal dispositions" (p. 80). There
seems to be a corresponding tendency to believe that the observers' (assessors') perceptions are objective and accurate, while the perceptions of those assessed (actors) are subjective and biased.

Systematic observer–actor differences have been interpreted as arising from (a) information availability and processing, and (b) motivation—including self-serving interpretations (Bradley, 1978; Miller & Ross, 1975; Monson & Snyder, 1977). Research has found observer–actor differences consistently enough to warrant speculation that this might be an important biasing factor in assessment. For example, a great many assessors may be predisposed to look for, and localize, causes of psychoeducational problems within those who manifest the problems, the students. In contrast, students may tend to attribute cause to teachers, peers, or task difficulty. If this is to be expected as a result of psychological influences, how can one know whose perceptions are correct? Perhaps both are biased and incorrect. Selectivity and distortion in assessment that results from actor–observer differences in perception clearly needs further investigation (Compas, Adelman, Freundl, Nelson, & Taylor, 1982; Compas & Adelman, 1981).

Professional role demands are also potential factors causing selectivity and distortion. Psychologists and educators are under constant pressure to demonstrate greater competence than current knowledge warrants. This pressure often leads professionals to overstate expertise. In turn, their public declarations then lead them to defend, and subsequently believe in, these unsubstantiated theories and practices. Other individuals oversell their expertise, not for ego-defensive reasons, but for ego-enhancing purposes, including such self-serving objectives as attaining status and financial rewards.

As these examples suggest, the cognitive–perceptual–motivational predispositions of assessors can combine in complex ways to shape their actions and conclusions.

In the case of David, the teacher, psychologist, and his parents started out with the assumption that something was wrong with him. Thus it is not surprising that what they looked for and found were problems within David. Then, it seemed only logical to help him deal with his problems.

**Societal bias.** Prevailing social values define what is exceptional, deviant, or a problem, and how it should be dealt with. These values are translated into governmental policies that determine which psychoeducational problems are attended to and which practices are encouraged. As many writers have suggested, interventions are designed as much—and perhaps more—to serve and protect society's interests as they are to help an individual. In particular, it has been argued that the best developed assessment practices, especially widely used tests, are primarily a reflection of society's need to foster and protect its own interests.

Chase (1977) presents the position that society, for self-serving purposes, has perpetuated biological explanations for problems that require social solutions. Kamin (1974) proposes that IQ tests have been used in ways that support "the belief that those on the bottom are genetically inferior victims of their own immutable defects." He further suggests:
The courts have ruled that some testing procedures result in racially discriminating practices.
the consequence has been that the IQ test has served as an instrument of oppression against the poor—dressed in trappings of science, rather than politics. . . . The poor, the foreign born, and racial minorities are shown to be stupid [and] born that way. The underprivileged are today demonstrated to be ineducable, a message as soothing to the public purses as to the public conscience. (p. 2)

Similarly, Coles (1978) argues that the “biologizing” of social problems has resulted in the positing of organic causalities for poverty, aggression, and violence, as well as for educational underachievement. . . . By positing biological bases for learning problems, the responsibility for failure is taken from the schools, communities, and other institutions and is put squarely on the back, or rather within the head, of the child. Thus, the classification (assessment leading to the diagnosis of learning disabilities) plays its political role, moving the focus away from the general educational process, away from the need to change institutions, away from the need to rectify social conditions affecting the child, and away from the need to appropriate more resources for social use toward the remedy of a purely medical problem . . . a classic instance of what Ryan (1971) has called “blaming the victim.” (p. 333)

What is being boldly stated by such writers is that society has a large stake in how the causes and corrections of psychoeducational problems are understood. Bias stemming from society’s values seems so ubiquitous a phenomenon that some-
one (source unknown) has formulated the “law of selective attention to data.” This “law” postulates that the greater the ideological relevance of research or assessment findings, the greater the likelihood that involved professionals will selectively pay attention to the data gathered.

Prevailing biases are reflected in mandatory programs and in research and development activity supported by government. Because societal bias so pervasively shapes perceptions, it is likely that most people are unaware of the selectivity and distortion built into government supported practices. This lack of awareness, of course, makes it difficult to counter the bias.

In the case of David, a teaching assistant in the classroom had noticed that two other children seemed equally inattentive and were falling behind in their work. The others, both girls, stared quietly at their books and drew little attention to themselves. David, on the other hand, got a bit noisy when bored and tapped his pencil, flipped pages, and talked to others at his table. Since she was concerned with sex role discrimination, the assistant wondered whether noisy boys were more likely than quiet girls to be referred for testing and special programs. More generally, she wondered whether such a societal bias might be a significant factor in learning disabilities being seen as primarily a boy’s problem (see Feature 4-4).

In general, there seems to be a widespread impression that existing assessment procedures are better than they are. This impression contributes to an

**Feature 4-4**

**Why Are More Boys than Girls Diagnosed as Having Learning Disabilities?**

It is important to keep in mind the full range of possible causes, rather than to assume that most learning problems are related to neurological troubles. Only from a broad perspective can one appreciate the difficulty of adequately explaining such facts as the overall higher incidence of learning problems among males (and low-income minority groups) and the underachieving patterns that appear among females in their later schooling. For example, available data indicate that 72 percent of those diagnosed as having learning disabilities are boys (U.S. General Accounting Office, 1981; also see Finucci & Childs, 1981). Those who believe current diagnoses of learning disabilities are accurate tend to explain the sex differential in terms of biological differences.

One set of theories stresses sex differences in neurological development and specialization of the cerebral hemispheres (Dalby, 1979). Another explanation suggests that, on the average, males have larger heads at birth than females, and this greater size increases the probability of difficult births and higher rates of oxygen deprivation causing brain damage. In contrast, those who believe that a significant number of males are misdiagnosed as having learning disabilities tend to look for explanations outside the central nervous system. They hypothesize, for instance, that the source of many males’ learning problems and of the underachievement patterns of many girls in their later schooling can be traced to differences in cultural expectations. Many boys come to school with experience that tends to make them more interested in highly active pursuits than in quiet academic activity. As they get older, many girls are expected to behave in ways that play down their academic capabilities. Comparable points about psychological, socioeconomic, cultural, and political factors have been made about the high incidence of learning problems among children reared in poverty.
unquestioning use of certain procedures in making major psychoeducational decisions. Methodological deficiencies (in terms of reliability, validity, and norms) should raise sufficient concerns about such practices. Moreover, the weak knowledge base underscores concerns about (1) what is actually being assessed, (2) overreliance on tests, (3) the low utility of data gathered with respect to decision making, and (4) what factors actually play the most significant role in arriving at a particular type of decision. In this section, we have also stressed that current procedures may be biased. Awareness of such bias underscores that, in many instances, it is the assessor’s rationale (rather than valid findings) that guides and shapes assessment decisions. Obviously, work is needed to improve assessment procedures and counteract the problems of selectivity and distortion that permeate current assessment practices.

**Ethics**

Impetus for ethical and related policy concerns about negative aspects of assessment has come from reported misuses and abuses of test data. Criticism has come from political conservatives, liberals, and civil libertarians. At the center of the controversy is the traditional tension between society’s rights, responsibilities, and needs, and individuals’ rights and freedoms. Critics have argued that individual rights and liberties are not sufficiently safeguarded and have pushed for greater legal protection of rights and due process. This is leading to improvements in consent procedures.

Another line of ethical criticism and policy concern stresses the errors, costs, and negative side effects of assessment. Some critics cite psychological, social, economic, and possible physical harm to individuals; others point out that subgroups are discriminated against; a few have raised the specter of quality of life in society may be significantly lowered by institutionalizing assessment practices. In contrast, some professionals underscore that it is an ethical responsibility of professionals to use assessment practices to maximize benefits for individuals and society.

The widespread use of assessment to define learning problems is ample indication of belief in its benefits. In this section, however, we highlight the two most prominent areas of ethical concern—privacy rights and negative consequences of assessment.

**Privacy.** There is a dual concern about privacy rights: invasion of privacy and misuse of information. These concerns arise when the information is considered highly sensitive and could lead to diagnoses and evaluations that are perceived negatively. The situation is especially volatile when assessment is carried out primarily to serve societal or institutional objectives.

Power to assess—to obtain and use information about others—is power to shape lives. Legally and ethically, there is a need to keep such power in check. At issue is the nature of the control people ought to have over the gathering and disclosure of information about themselves. In other words, when should society be able to mandate assessment if, in the process, it infringes on individual rights? This is an aspect of the broader concern over when society should be allowed to
coerce individuals and, thereby, deny a variety of rights and freedoms (Adelman & Taylor, 1988; Robinson, 1974). From this issue springs a variety of questions. What kind of information is it reasonable to gather on an individual? What safeguards exist with respect to the collection and use of highly personal and sensitive information? What types of records should be kept and who should have access to them? What restrictions should be placed on how information can be used? Is parental consent sufficient when children don’t want to be assessed?

The complexity of ethical concerns is well illustrated when individuals come for help. A request for help may be seen as consent to gather data on anything the assessor sees as relevant. Given adequate theory and evidence about what is relevant, it would be a relatively clear-cut matter to explain what is needed and why as a basis for eliciting informed consent. Unfortunately, the state of knowledge regarding psychoeducational problems is not sophisticated enough to specify absolutely what information is needed. Thus assessors develop their own criteria. Some feel free to follow their intuition about anything that seems significant at the moment. Although well meaning, they may pry into painful areas of a person’s life to gather data that may be irrelevant. Such data may amount to little more than gossip. Insignificant and invalid data can be even more harmful when used inappropriately (for example, to prescribe treatment).

For a variety of political and legal reasons, many school systems have moved away from presumptions of consent. In the United States, this movement has been accelerated by federal law (for example, the Family Educational Rights and Privacy Act of 1974). The apparent result has been that (1) less assessment data are gathered and circulated in schools, (2) consent is sought more frequently when a need for assessment exists, and (3) due process is emphasized with respect to student and parent access to records and for complaints and corrections of data that may be inappropriate or in error.

Critics caution, however, that the burden of protecting rights still falls mostly on those assessed. Consumer advocates advise students or parents to object if they dislike either what is asked or the procedures. When they do so, however, they run the risk of being refused services or having the objection interpreted as defensiveness, hostility, or lack of cooperation.

The situation is further complicated where procedures have become routinized and institutionalized. Under such circumstances, those involved may see neither a rights issue nor a need for consent. For example, physicians, psychologists, educators, and a variety of other personnel in schools, clinics, and agencies routinely administer tests and questionnaires with little or no explanation about why the information is needed or about any limitations with respect to the procedures’ validity. When procedures have become a natural part of an institution’s operations, those administering them may be genuinely unaware of invading privacy or coercing people. Those who are assessed may assume that the experience must be essential, and any discomfort is to be borne silently.

**Negative consequences.** As stressed in Chapter 3, every major intervention has some negative consequence. Assessment is no exception. It is customary to speak of “negative side effects,” but this wording tends to ignore errors and
economic costs and is more appropriately applied to minor and perhaps low-probability phenomena. Negative consequences encompass the range of potentially significant harm that may occur.

Negative consequences, such as extreme anxiety, may occur during the assessment process or may be an immediate or long-term outcome. It is widely recognized that persons who are assessed and labeled may be stigmatized, isolated, and excluded from important experiences, and that this may negatively affect motivation and hinder full and healthy development. Evidence suggests that certain subgroups are more likely than others to experience such negative effects (Copeland, 1983; Heller et al., 1982; MacMillan, Hendicks, & Watkins, 1988; Swanson & Watson, 1982). Civil rights hearings and court cases have highlighted the intentional and unintentional cultural and sex-role bias of formal assessment (see Feature 4-5). Unfortunately, little data exist on the frequency of negative consequences (including inevitable errors) or about financial costs to individuals and to taxpayers. Concerns about costs are increasing. For example:

In the case of testing to identify children in mildly handicapped categories, the costs of assessment and staffing procedures use up half of the extra per-pupil resources available without any evidence that pro forma administration of tests adds to the scientific integrity of placement decisions.
Feature 4-5  The Changing Sociopolitical Climate

Throughout the 1970s and 1980s, judicial decisions were an important sociopolitical factor influencing legislation, policy making, and practice. In the courts, parents and advocacy groups pushed for the rights of those with disabilities and equality of education and fought discrimination (for example, see Litigation and special education, 1986).

Hobson v. Hansen (1969)  This was the first major case raising questions about placement in special education. The court ruled that using test scores to group students into “tracks” was unconstitutional because it discriminated against blacks and the poor.

Diana v. State Board of Education (1970)  Here was a case in which the use of tests to place students was again challenged. Diana, a Spanish-speaking student in Monterey County, California, had been placed in a class for mildly mentally retarded students because she had scored low on an IQ test given to her in English. The court ruled that Spanish-speaking children should be retested in their native language to avoid errors in placement.

Mills v. D.C. Board of Education (1972)  Seven school-age students had been excluded from school because of handicaps. The school district claimed it did not have the money needed to provide an appropriate education for them. The landmark decision in this case established that children with disabilities should be given a free and suitable public education and cannot be excluded for financial reasons.

Pennsylvania Association for Retarded Citizens (PARC) v. Commonwealth of Pennsylvania (1972)  Like the Mills case, this was a class-action suit that declared that all children have the right to a free and appropriate education.

Larry P. v. Riles (1972)  Larry P. was a black student in California, and his complaint led to an expansion of the ruling in the Diana case. The court ruled that schools are responsible for providing tests that do not discriminate on the basis of race. In the class-action case of PASE v. Hannon (1980), however, the judge stated he could find little evidence of bias in the test items. The Larry P. case also set a precedent for the use of data indicating disproportionate placement of minority groups as prima facie evidence of discrimination. However, subsequent cases have undermined this precedent (Marshall et al. v. Georgia [1984] and S-I v. Turlington [1986]).

Le Banks v. Spears (1976)  The decision in this case helped spell out the nature of a suitable or appropriate education for those with disabilities.

Burlington School Committee v. Department of Education (1985)  The U.S. Supreme Court ruled that parents who place their disabled child unilaterally into a private school are entitled to tuition reimbursement after a court determines the placement was appropriate under Public Law 94-142 and the public school placement was inappropriate.

Honig v. Doe (1988)  The U.S. Supreme Court ruled that Public Law 94-142 prohibits schools from unilaterally excluding students with disabilities from the classroom for dangerous or disruptive conduct growing out of their disabilities. Normal nonplacement-changing procedures, including temporary suspensions for ten school days, are allowed. This case appears to overrule the case of Victoria L. v. District School Board (1984), which did allow for unilateral placement of a dangerous disabled student pending review proceedings.

In many ways, it is unfortunate that litigation is necessary. But it is. Current legislation may mandate rights, but as Meyen (1982) has stated, it "does not guarantee that compliance will be enforced" (p. 9). Moreover, it seems that many questions about the appropriate treatment of those who do not "fit in" are so unsettled and unsettling as to require adjudication.
[Research shows] there is a very high correspondence between initial teacher referrals and final placement decisions, with all the testing in between serving to justify placement. At least half of the children labeled by schools as learning disabled [LD], by far the largest category of handicap, are misidentified. Rather than fitting the original definition of LD, they are more aptly described as slow learners, linguistically different children, misbehaving boys, children who are absent or whose families move too frequently, or as average learners in above-average contexts. (emphasis added, Shepard, 1990: p. 23)

Illich (1977) and others have warned that the cost to society may be more than financial. He has argued that overreliance on professionals leads to an alarming incapacity among individuals and natural support systems to cope with problems. The rapid rise in number of children diagnosed as having learning disabilities, and the highly specialized (and sometimes inappropriate) treatments prescribed may be a most poignant example of such effects.

From a practical perspective, concern over negative consequences generally centers on how to minimize negative effects and be certain that benefits outweigh harm. Often at issue is whether the positive outweighs the negative. There does seem to be widespread agreement, however, about the following guidelines:

1. Assessors are obligated at least to be aware of potential negative consequences, such as immediate and long-term harm to individuals, groups, and society.
2. Where consent is sought, assessors are required ethically and often legally to inform prospective consenters of potential positive and negative consequences.
3. As they attempt to maximize benefits, assessors are obligated to minimize potential negative effects.
4. Assessors are expected not only to look for data that confirm hypotheses about problems, they also must actively seek disconfirming evidence.
5. Although they cannot follow a student around to prevent self-fulfilling prophecies, they are expected to take steps to correct and guard records and equip students and parents to protect and advocate for themselves.
6. Assessors are expected to acknowledge whenever findings are inconclusive and not rationalize or dismiss uncertainties and incongruities in findings.

**Meeting Minimal Standards**

Although critics have consistently raised concerns, there are few who would argue that a procedure should not be used until it is perfected (if it ever could be). The consensus is that, despite inadequacies, many assessment tools can provide useful information to inform decision making—if they are properly chosen and used. But what constitutes proper choice and use?

At the very least, most will agree that assessment procedures should meet the minimal standards set forth by the American Psychological Association and the American Educational Research Association (see "Standards for Educational and Psychological Tests"). From an ethical and legal perspective, a practice should do more good than harm. In this context, it has been argued that some psychometric
and rating-scale procedures used for massive screening are used in ways that produce so many false-positive errors that they probably cause more harm than good and should not be used (Adelman, 1989). In California, based on the Larry P. v. Riles court decision, the Los Angeles Unified School District has taken the extreme position of placing a moratorium on the use of IQ test data in special education placement decisions.

When assessment procedures are seen as meeting minimal standards and providing useful information, there is still the problem that data gathered may be limited and perhaps erroneous. There are no satisfactory solutions for these problems. A common suggestion is that additional data be gathered that might at least disconfirm false-positive findings. Ethically, it has been stressed that all concerned parties should be alerted to the limitations of the findings and the tentative nature of recommendations and conclusions based on the data. With respect to gathering disconfirming data, there is growing advocacy for viewing decision making as a multi- rather than single-stage process.

| Beyond Conventional Practices |

In this section, we explore four directions for change advocated by those who have been critical of conventional practices.

**Preassessment Interventions**

Critics concerned about premature person-focused assessment have argued that major efforts to improve programs should come first. In this context, we recall Hobb’s (1975a) views on screening.

Ideally, special screening programs to identify health problems and developmental difficulties of children should not be necessary. All children regardless of economic status should be able to participate in a comprehensive health maintenance program. (pp. 90–91)

We would add that, once children arrive in kindergarten and the elementary grades, they should be provided with comprehensive psychoeducationally oriented school programs. That is, the need for screening is secondary to the need to evolve classroom programs to match learners’ levels of motivation and development. Advocates suggest that preventive and early-age intervention programs can reduce the number of learning and behavior problems, provide improved in situ screening to identify those who continue to require special help, and thereby reduce the need for special assessment.

As indicated in preceding chapters, there are a substantial number of advocates for improving programs as the first step in a screening sequence for learning problems (see Feature 4-6). In essence, they argue that, prior to assessment, programs to ameliorate learning problems should (a) enhance regular learning and instruction for all children, and (b) remedy problems that arise as soon as possible and with the least intervention needed. It is recognized that accomplishment of these goals requires broadening the psychoeducational nature of regular school
interventions (personalizing instruction) and increasing the availability of support mechanisms for academic learning (volunteer aides, peer tutoring, computers). Proponents also stress the need for programs to train and provide consultation and peer collaboration opportunities for teachers regarding what interventions to pursue before referring a student for special education (prereferral interventions).

**Focal Point**

There is general concern that the focus of prevailing assessment procedures encompasses too narrow a range of factors. This criticism is being voiced more strongly as new models emerge; these models challenge the prevailing view that sees learning problems in terms of person variables (that is, in terms of pathology or lack of readiness). Competing models focus on the environment (also emphasizing either pathology or deficiencies) or the transactions between person and environment.

Of the three models, only the person model has been used extensively. Its focus has been on assessing physiological and psychological correlates of pathology, developmental deficits, or both. But, as can be seen in texts on special education and neuropsychological assessment, this focus has been restricted to a highly delimited set of variables associated with behavior and learning (Gaddes, 1985; Salvia & Ysseldyke, 1991; Swanson & Watson, 1982). Some researchers have suggested broadening the range of assessed correlates to include social skills, temperament and cognitive style, and a variety of cognitive and metacognitive variables (Swanson, 1991). The focus of remedial planning and evaluation has broadened to encompass such variables. And there is increasing pressure to counter trends toward "skill and drill" assessment and teaching so as to avoid narrowing the curriculum and to foster higher-order thinking. One such trend has been called authentic or performance assessment.

Because of dissatisfaction with the prevailing person-oriented model, the work

---

**Feature 4-6 Proposed Research on Preassessment Intervention**

| Despite ongoing advocacy for improving programs as a first step in screening (Johnson & Pugach, 1991), the idea remains relatively uninvestigated. One type of study needed would upgrade a representative sample of preschool, kindergarten, and primary school programs across socioeconomic groups to improve personalized and remedial instruction. The proportion of children subsequently found to manifest problems in these settings would then be compared with those in a matched control sample of standard programs. Data from this comparison would indicate the efficacy of the experimental settings in preventing some types of problems. | Identification of learning problems in each classroom would involve no more than establishing criteria for daily performance and noting those who do not meet the criteria over a period of several weeks. Students in the experimental and control samples could be followed into the upper elementary grades to determine the degree to which false-positive and false-negative identifications were made. After this first study, a second could determine whether identification sensitivity and specificity are improved by adding formal assessment screening procedures to the experimental programs. |

---
of researchers who focus on the environment has taken on some prominence. These investigators assess home and school variables to clarify the role they play in learning problems in general and learning disabilities in particular (Barclay, 1983; Fraser & Walberg, 1991; Freund, Bradley, & Caldwell, 1979; Lloyd & Blandford, 1991; Moos, 1979).

Going a step further, transaction-oriented investigators hope to determine the degree to which the interplay of person and environment must be accounted for in understanding cause and correction (Adelman, 1970–71; Bandura, 1978; Coles, 1987; Sameroff, 1985). Those who have adopted a transactional (interactional, reciprocal determinist) orientation have incorporated, not rejected, the other two models.

**Single- versus Multi-Stage Decision Making**

Although some assessors find it necessary or convenient to assess and make decisions in one or two sessions, such a single-stage approach is a matter of concern. For example, it is generally acknowledged that, after a potential problem has been identified, subsequent steps must be taken to confirm or disconfirm it. (Critics warn that diagnoses and placement decisions often are made solely on the basis of first-level screening data.) Furthermore, even when the best available assessment procedures are used, initial decisions about placement and special programming may be in error and should be confirmed or disconfirmed through monitoring performance as the individual pursues daily tasks.

Similar arguments have been raised about ways to improve diagnosing learning disabilities and learning disability subgroups. For instance, researchers suggest that identifying those whose learning problems are due to CNS trouble requires filtering them out through sequential assessment (Adelman, 1971; Lindsay & Wedell, 1982; Wissink, Kass, & Ferrell, 1975). Increasingly, sequential or multi-stage assessments are advocated as one way to work on improving decision accuracy, including reducing the number of persons wrongly diagnosed as having learning disabilities.

**Interventionist Assessment**

Conventional, psychometric approaches and techniques raise a variety of validity concerns. A common example already noted is that, under formal assessment conditions, poor performance among problem populations may be due to low or negative (avoidance) motivation resulting from high anxiety or negative attitudes. These factors, and cultural differences, may negatively affect the performance of persons from certain cultural backgrounds. That is, the assessment results for such persons are seen as contaminated and thus cannot be taken at face value. These critics argue that it is impossible to know whether failure to demonstrate specific knowledge, abilities, or skills represents a real deficiency. The implications for research and intervention are profound.

Within the psychometric tradition, efforts to deal with this criticism stress accounting for contaminants when interpreting findings, improving task content and administration to reduce biasing conditions (including frequent reassess-
ments), and going beyond standardized administration to assess how much more the individual can do (for example, allowing additional time, adding a brief prompting or teaching facet to the process). Going a step further, those in the dynamic assessment movement argue for alternative procedures designed to determine how much more the individual can do when comprehensively prompted and taught. This is a highly interventionistic approach to assessment. The assessor is "an active intervener who monitors and modifies the interaction with the learner in order to induce successful learning. The learner is prodded, directed, and reinforced into a role of active seeker and organizer of information" (Lidz, 1987; pp. 3–4).

To underscore the fundamental direction involved here, it is useful to contrast conventional practices with what can be described broadly as interventionist assessment. The term interventionist goes beyond dynamic assessment to encompass a wide variety of activities designed to determine whether a person can perform at a higher level. Interventionist assessment is designed to move beyond the nondynamic/static approach found in typical psychometric testing. The assessor actively "tests the limits" to encourage increased performance. This often takes the form of an assess–teach–reassess approach, involving a reasonable interval of time for learning to take place.

Critics raise questions about the underlying assumptions and the validity of interpretations made by interventionist assessment. For example, they question whether deficient functions found in cognitive processes can be modified to a significant extent during an assessment procedure. With respect to the validity of interpretations, they also charge that performance changes may primarily reflect motivational rather than cognitive change.

Advocates of interventionist assessment approaches state that the intent is not to replace but to supplement prevailing assessment procedures (Lidz, 1987; Palincsar, Brown, & Campione, 1991). They stress they are seeking data not available through prevailing approaches (for example, data on performance capability in a teaching situation, and information on teaching approaches that appear to be effective with the learner).

---

**Summing Up**

Assessment is a broad-based concept. The term has been adopted to encompass narrower, medically related processes such as diagnosis, screening, and diagnostic testing. Formally defined, assessment is the process by which attributes of phenomena are described and judged. Descriptions take the form of data gathered by formal and informal measures, such as tests and observations of behavior or settings. Judgments take the form of interpretive conclusions about the meaning of data.

Psychoeducational assessment encompasses four major purposes:

1. identification (including screening)
2. selection (including placement)
3. planning for specific change
4. evaluation of intervention

Choices about what data to gather and what to exclude are guided by the types of judgments and decisions to be made (diagnostic classification, placement, remediation).

Controversy surrounds prevailing approaches to assessing learning problems. Although some of the controversy is about the deficiencies and limitations of specific procedures, broader concerns and criticism have been directed at the way assessment is used to shape research and practice and related policy decisions. Even when relatively objective assessment data are used, subsequent decisions often are extremely subjective. This is not surprising, given that most decisions involve considerations that go well beyond the availability of valid data. More often than not, complex social–political–economic value questions are involved. Indeed, in some cases seemingly relevant data are ignored in order to arrive at a decision that the decision makers see as viable and beneficial (Woodhead, 1988). Thus controversy is inevitable, and as Thorndike and Hagen (1977) have aptly stated, “The wisdom of the decider is crucial” (p. 20).

What should be clear by this point is that assessment for identification, placement, program planning, and evaluation is a complex matter. Moreover, the state of the art is seriously restricted. Thus, despite the importance of all four assessment functions, prevailing assessment procedures

- do not have sufficient validity to warrant large-scale programs aimed at early identification;
- are not capable of producing appropriate differential diagnoses and placements for persons with learning problems;
- misprescribe remediation, deemphasize the importance in program planning of a person’s strengths and interests, and narrow the focus of school curricula; and
- inaccurately shape evaluation and eventually redefine and limit objectives.

Furthermore, overemphasis on assessment practices that focus on persons hinders development of procedures for assessing the role of the environment. As a result of the bias toward localizing problems within persons, interventions tend to be person-centered. Almost by presumption, environmental variables are exonerated as causal factors and as the focal point of intervention.

In spite of the deficiencies of prevailing practices, each day professionals are called upon to assess and make decisions about individuals with learning problems. Unfortunately, for now they must do so using a relatively weak knowledge base.

The need for improved practices is evident. Fortunately, recent research has pointed to promising approaches that go beyond conventional procedures. There may be major concerns about the state of the art, but there can be no doubt that persons with problems can and must be helped. And, because learning problems are a life-long concern, such help must be available to persons of all ages. Chapter 5 highlights program considerations from preschool to postschool.
I. Think about a test on which you received a low grade:

1. Was the test a reliable and valid way to find out what you had learned?
2. How do you explain the fact that you received a low grade?
3. What type of attributions do you think the grader made about your test performance?
4. What do you think is the best way to assess what you have learned?

II. Engage a group of friends in a discussion about the pros and cons of using labels such as learning disabilities and placing those with learning problems in special classes.

   Explore any concerns group members have about current assessment procedures that are used for diagnosis and placement.

   Is the use of college admission tests an example of using assessment for identification and selection?

---

**On Assessment Tools**

*Technical Concerns About Assessment* (p. 349) This contains a brief overview of basic concepts that are central to understanding how good an assessment procedure is. The four topics are:

- reliability
- validity
- norms
- standards

*Procedures and Instruments for Assessing Learning Problems* (p. 341) As a special resource, this (a) provides a list of specific assessment procedures and instruments frequently used in assessing psychoeducational problems, and (b) gives references for finding and evaluating them.

**On Learning Disabilities**

*Assessment for Learning Disabilities: Diagnosis, Placement, and Program Planning* (p. 328) With respect to diagnosing, placing, and planning programs for learning disabilities, this

- reviews assessment approaches, and
- highlights concerns about certain practices.

*Screening for Learning Disabilities* (p. 323) Searching for learning disabilities is both common and controversial. The discussion stresses

- the nature and scope of screening practices, and
- what research suggests about the state of the art.
Remedying Learning Disabilities: Prevailing Approaches (p. 314) This explores major approaches to remedying learning disabilities through contrasting orientations: It discusses (1) underlying-problem approaches, with their focus on perception, motor functioning, language, and general cognitive functioning, and (2) observable-problem approaches, with their focus on observable skills and objectives, and direct instruction.